PCT/GB 2004 / 0 0 2 7 0 6

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed R. Mahoney

Dated 8 July 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

150EC03 E859201-1 D00027
P01/7700 0.00-0328884.2 NONE

1/77

THE PATENT OFFICE
12 DEC 2003
LONDON

The Patent Office

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office
Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference 61.82412

2. Patent application number *(The Patent Office will fill in this part)* 0328884.2 12 DEC 2003

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Isis Innovation Limited
Ewert House
Ewert Place
Summertown
Oxford
OX2 7SG
United Kingdom

Patents ADP number *(if you know it)* 39985640003.

If the applicant is a corporate body, give country/state of its incorporation United Kingdom

4. Title of the invention Methods

5. Name of your agent *(if you have one)* Frank B. Dehn & Co.

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

179 Queen Victoria Street
London
EC4V 4EL

Patents ADP number *(if you know it)* 166001 ✓

6. Priority: Complete this section if you are declaring priority from one or more earlier patent applications, filed in the last 12 months

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. Divisionals, etc: Complete this section only if this application is a divisional application or resulted from an entitlement dispute *(see note f)*

Number of earlier UK application	Date of filing *(day / month / year)*

8. Is a Patents Form 7/77 (Statement of inventorship and of right to grant of a patent) required in support of this request?

Answer YES if:

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
Otherwise answer NO (See note d)

Yes

Patents Form 1/77

9. Accompanying documents: A patent application must include a description of the invention. Not counting duplicates, please enter the number of pages of each item accompanying this form:

Continuation sheets of this form	0
Description	35
Claim(s)	2
Abstract	0
Drawing(s)	0

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for a preliminary examination and search *(Patents Form 9/77)*	
Request for a substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. **I/We request the grant of a patent on the basis of this applicati**

Frank B. Dehn & Co.

Signature **Date** 12 December 2003

12. Name, daytime telephone number and e-mail address, if any, of person to contact in the United Kingdom

Annabel R. Beacham
01273 244200

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*
b) *Write your answers in capital letters using black ink or you may type them.*
c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s) of the form. Any continuation sheet should be attached to this form.*
d) *If you have answered 'Yes', Patents Form 7/77 will need to be filed.*
e) *Once you have filled in the form you must remember to sign and date it.*
f) *Part 7 should only be completed when a divisional application is being made under section 15(4), or when an application being made under section 8(3), 12(6) or 37(4) following an entitlement dispute. By completing part 7, you are requesting this application takes the same filing date as an earlier UK application. If you want the new application to have the same priority date(s) as the earlier UK application, you should also complete part 6 with the priority details.*

82412.209

Methods

The present application is concerned with methods for the chemical modification of proteins, in particular methods for the glycosylation of proteins.

The co- and post-translational glycosylation of proteins plays a vital role in their biological behaviour and stability (R. Dwek, Chem. Rev., 96:683-720 (1996)). For example, glycosylation plays a major role in essential biological processes such as cell signalling and regulation, development and immunity. The study of these events is made difficult by the fact that glycoproteins occur naturally as mixtures of so-called glycoforms that possess the same peptide backbone but differ in both the nature and the site of glycosylation. Furthermore, since protein glycosylation is not under direct genetic control, the expression of therapeutic glycoproteins in mammalian cell culture leads to heterogeneous mixtures of glycoforms. The ability to synthesise homogeneous glycoprotein glycoforms is therefore not only a prerequisite for accurate investigation purposes, but is of increasing importance when preparing therapeutic glycoproteins, which are currently marketed as multi-glycoform mixtures (e.g. erythropoietin and interleukins).

A number of methods for the glycosylation of proteins are known, including chemical synthesis. Chemical synthesis of glycoproteins offers certain advantages, not least the possibility of access to pure glycoprotein glycoforms. One known synthetic method utilises thiol-selective carbohydrate reagents, glycosylmethane thiosulfonate reagents (glyco-MTS). Such glycosylmethane thiosulfonate reagents react with thiol groups in a protein to introduce a glycosyl residue linked to the protein via a disulfide bond (see for example WO00/01712).

However, glyco-MTS reagents suffer from a number of disadvantages, including occasionally moderate reaction

yields, difficulties in their preparation and problems with stability under the basic conditions in which they are often used.

There is therefore a need for alternative methods for protein glycosylation which give high yields of the glycosylated protein product, are site-selective, and which allow glycosylation at both single and multiple sites in a wide range of different proteins.

We have now surprisingly found that a thiol containing protein may be converted to the corresponding selenenylsulfide, and that the electrophilic character of the sulfur in the S-Se bond thus created renders it susceptible to nucleophilic substitution by thiol-containing compounds including carbohydrates.

In one aspect, the invention therefore provides a method of chemically modifying a protein, peptide or amino acid comprising at least one thiol group (-S-H), the method comprising converting said thiol group into a selenenylsulfide group (-S-Se-R). The method therefore allows the preparation of a protein, peptide or amino acid comprising at least one selenenylsulfide group. Such proteins, peptides and amino acids comprising at least one selenenylsulfide group form a further feature of the invention. Particularly preferred are proteins or peptides comprising at least one selenenylsulfide group.

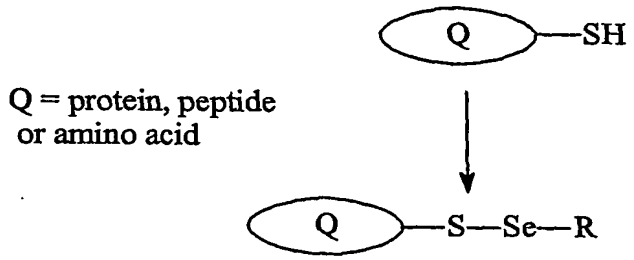
A selenenylsulfide group in a protein, peptide or amino acid may be further reacted with an organic compound comprising a thiol group to give further chemically modified proteins, peptides or amino acids in which the organic group is attached to the protein, peptide or amino acid via a disulfide bond. Preferably, the organic compound containing the thiol group is a carbohydrate compound, thus providing a method for the glycosylation of an amino acid, peptide or protein. As used herein, "glycosylation" refers to the general process of addition of a glycosyl unit to another moiety via a covalent linkage.

In a further aspect, the invention therefore provides a method of chemically modifying a protein, peptide or amino acid comprising at least one thiol group (-S-H), the method comprising:

(a) converting said thiol group into a selenenylsulfide group (-S-Se-R); and

(b) reacting said selenenylsulfide group with an organic compound containing a thiol group.

A generalised reaction scheme for the introduction of a selenenylsulfide group into a protein, peptide or amino acid is shown in Scheme 1:

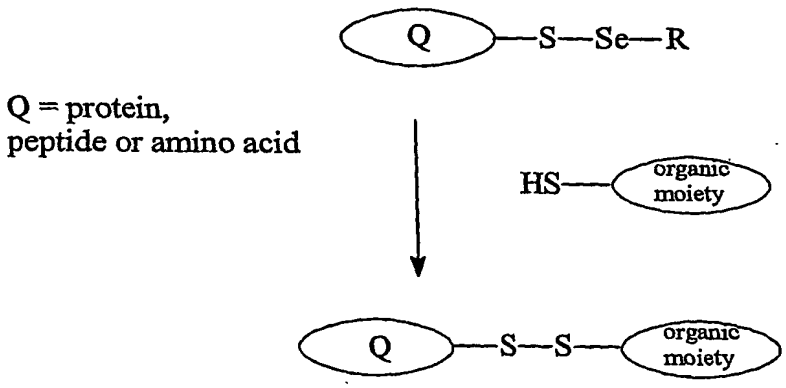


Scheme 1

The method of Scheme 1 results in covalent linkage of a group R to the protein, peptide or amino acid via a selenenylsulfide (-S-Se-) linkage. Such proteins, peptides or amino acids form a further feature of the invention. Proteins and peptides comprising a selenenylsulfide group may be useful in the determination of protein structure via X-ray diffraction techniques. Currently, MAD (multiple wavelength anomalous dispersion) techniques involve the conversion of any methionine residues in the protein into selenomethionine. Comparison of the X-ray diffraction patterns of the modified and unmodified proteins then allows a determination of the structure of the unmodified protein to be carried out. The method of the invention allows convenient and ready access to

alternative selenium-containing proteins or peptides which may be used in such techniques.

Such selenenylsulfide containing proteins, peptides or amino acids may be further reacted with thiol containing organic compounds as shown in the generalised reaction scheme in Scheme 2:



Scheme 2

The method of Scheme 2 results in covalent linkage of the organic moiety to the protein, peptide or amino acid via a disulfide bond (-S-S-). In this method the protein, peptide or amino acid is acting as an electrophile whilst the thiol-containing organic compound acts as a nucleophile. In contrast, the known reactions utilising glyco-MTS reagents involve reaction of a nucleophilic thiol group in the protein, peptide or amino acid with the electrophilic glyco-MTS reagent. The method of the invention therefore provides a complementary strategy to the known protein modification strategies utilising glyco-MTS reagents.

The conversion of the at least one thiol group in the protein, peptide or amino acid to a selenenylsulfide group is highly selective. In addition, the reaction of the thiol containing organic compound with the selenenylsulfide group is highly site-selective. It is not therefore normally necessary for any other functional groups in the protein, peptide or amino acid

or in the thiol containing organic compound to be protected whilst practising the methods of the invention. This can be highly advantageous, as it avoids the need for any subsequent deprotection steps to be carried out on the product.

If the protein, peptide or amino acid comprises more than one thiol group, then each such thiol group will potentially be converted to the corresponding selenenylsulfide group. Each such selenenylsulfide group may then potentially be reacted with a thiol containing organic compound, leading to attachment of the organic compound via a disulphide linkage to the protein, peptide or amino acid at multiple sites. The methods of the invention therefore provides a convenient method for the chemical modification of a protein, peptide or amino acid at multiple sites. In particular, the methods of the invention allows glycosylation of a protein, peptide or amino acid at multiple sites.

As used herein, a peptide contains a minimum of two amino acid residues linked together via an amide bond. Any amino acid comprised in the protein, peptide or amino acid comprising at least one thiol group is preferably an α-amino acid. Any amino acid may be in the D- or L-form, preferably the L-form. The amino acid, peptide or protein may be any naturally-occurring amino acid, peptide or protein which comprises a thiol group, for example due to the presence of one or more cysteine residues. Alternatively, the amino acid, peptide or protein may be prepared by chemical modification of a precursor non-thiol containing amino acid, peptide or protein. Alternatively, a thiol containing peptide or protein may be prepared via site-directed mutagenesis to introduce a cysteine residue. Site-directed mutagenesis is a known technique in the art (see for example WO00/01712 and J. Sambrook *et al*, Molecular Cloning: A Laboratory Manual, 3rd Edition, Cold Springs Harbour Laboratory Press, 2001,

the disclosures of which are hereby incorporated by reference).

Preferred proteins include enzymes, the selectivity of which may be modified by controlled glycosylation using the methods according to the invention, and therapeutic proteins. Other preferred proteins include serum albumins and other blood proteins, hormones, interferons, receptors, antibodies, interleukins and erythropoietin.

Conversion of the thiol group in the protein, peptide or amino acid to a selenenylsulfide group is conveniently carried out by reacting said protein, peptide or amino acid with a compound of formula Ia or Ib:

$$\underset{\text{Ia}}{R\text{-}Se\text{-}L} \quad \text{or} \quad \underset{\text{Ib}}{R\text{-}Se(OH)_2}$$

wherein:

L denotes a leaving group, for example OH, Br, CN, Cl or I, preferably Br; and

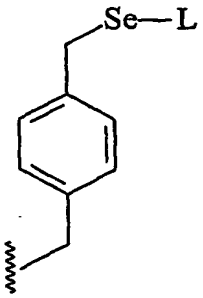
R denotes an optionally substituted alkyl group, an optionally substituted phenyl group, an optionally substituted benzyl group, an optionally substituted pyridyl group or an optionally substituted naphthyl group. A preferred R group is phenyl, a preferred compound of formula Ia is PhSeBr and a preferred compound of formula Ib is $PhSe(OH)_2$.

As used herein, alkyl preferably denotes a straight chain or branched alkyl group containing 1-10 carbon atoms, preferably 1-6 carbon atoms. Preferred alkyl groups include methyl and ethyl.

When R denotes an optionally substituted moiety, suitable substituents include any substituents which do not interfere with the reaction with the thiol containing protein, peptide or amino acid, and preferably also do not interfere with any subsequent reaction of the protein peptide or amino acid , for example reaction with a thiol containing organic

compound. Suitable substituents include $-NO_2$, $-SO_3H$, $-CO_2H$, $-(CH_2CH_2O)_nH$, and $-(CH_2CH_2O)_nMe$ wherein n denotes 1-100, preferably 1-50, more preferably 1-20, and still more preferably 1-10. The R group may be independently substituted by 1-5, and preferably 1 or 2, substituents.

The R group may also optionally be attached to, or form part of, a solid support. For example, the compound of formula Ia or Ib may be derived from a resin such as a polystyrene resin, as shown below:



The compounds of formula Ia and Ib are commercially available (e.g. PhSeBr, PhSeCl, PhSeCN, 2-nitrophenyl selenocyanate) or may be prepared by methods known in the art. For example, MeSeBr may be prepared according to the method of Hope, Eric G.; Kemmitt, Tim; and Levason, William, in Journal of the Chemical Society, Perkin Transactions 2: Physical Organic Chemistry (1972-1999) (1987), (4), 487-90, the disclosure of which is hereby incorporated by reference.

At least one mol equivalent of the compound of formula Ia or Ib per thiol group in the protein, peptide or amino acid should be used, to ensure conversion of each such thiol group to the corresponding selenenylsulfide group. The reaction is preferably carried out in an aqueous solvent (such as a mixture of water and acetonitrile) in the presence of a buffer (for example MES, pH 9.5). The pH and temperature of the reaction should be chosen such that undesirable denaturation of the protein or peptide is avoided. Preferably, the reaction is carried out at room

temperature or below, at a slightly basic pH (e.g. about pH 8 to about pH 9.5).

The organic compound containing a thiol group may be any organic compound which is suitable for linkage to a protein, peptide or amino acid, and in which the sulfur atom of the thiol group can act as a nucleophile to react with a selenenylsulfide group. There is no particular limitation on the nature of the organic compound. For example, the thiol group may be primary, secondary or tertiary. The compound may be aromatic or aliphatic. For example, the compound may be an alkyl, alkenyl or alkynyl thiol. Preferably, the compound only contains one thiol group.

Suitable organic moieties for attachment to a protein, peptide or amino acid include any group which may be useful in modifying the physical or chemical properties of the protein, peptide or amino acid. Suitable moieties include labels (for example fluorescent labels) or groups to aid the stability, processing or solubility of the protein, peptide or amino acid. The organic compound may also be a second protein, peptide or amino acid, giving the possibility of linking one protein, peptide or amino acid to another protein, peptide or amino acid via a disulphide linkage using the methods of the invention.

As used herein, alkenyl preferably denotes a straight chain or branched hydrocarbon group comprising at least one carbon-carbon double bond, and containing 2-10 carbon atoms, preferably 2-6 carbon atoms. Preferred alkenyl groups include $-(CH_2)CH=CH_2$ and $-CH_2CH_2CH=CH_2$. As used herein, alkynyl preferably denotes a straight chain or branched hydrocarbon group comprising at least one carbon-carbon triple bond, and containing 2-10 carbon atoms, preferably 2-6 carbon atoms. Preferred alkynyl groups include $-CH_2C\equiv CH$ and $-CH_2CH_2C\equiv CH$.

Preferably, the organic compound containing at least one thiol group is a carbohydrate moiety,

optionally attached via a linker to the thiol (-S-H) group. The linker may contain 1 to 10 atoms between the carbohydrate moiety and the -S-H group. For example, the linker may be an alkylene group (for example a $-(CH_2)_t-$ group wherein t denotes 1 to 10), or an alkenylene group (for example a $-(CH_2)CH=CH-$ or $-CH_2CH_2CH=CH-$ group).

Suitable carbohydrate moieties include monosaccharides, oligosaccharides and polysaccharides, and include any carbohydrate moiety which is present in naturally occurring glycoproteins or in biological systems. Preferred are optionally protected glycosyl or glycoside derivatives, for example optionally-protected glucosyl, glucoside, galactosyl or galactoside derivatives. Glycosyl and glycoside groups include both α and β groups. Suitable carbohydrate moieties include glucose, galactose, fucose, GlcNAc, GalNAc, sialic acid, and mannose, and oligosaccharides or polysaccharides comprising at least one glucose, galactose, fucose, GlcNAc, GalNAc, sialic acid, and/or mannose residue.

Preferred are compounds in which the thiol group is at the anomeric position of a saccharide residue or is attached to the anomeric carbon via a linker.

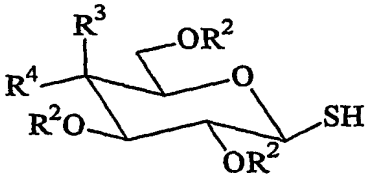
Any functional groups in the carbohydrate moiety may optionally be protected using protecting groups known in the art (see for example Greene et al, "Protecting groups in organic synthesis", 2nd Edition, Wiley, New York, 1991, the disclosure of which is hereby incorporated by reference). Suitable protecting groups for any -OH groups in the carbohydrate moiety include acetate (Ac), benzyl (Bn), silyl (for example tert-butyl dimethylsilyl (TBDMSi) and tert-butyldiphenylsilyl (TMDPSi)), acetals, ketals, and methoxymethyl (MOM). Any protecting groups may be removed before or after attachment of the carbohydrate moiety to the amino acid, peptide or protein. Preferably, they are removed before reaction with the selenenylsulfide compound, to remove the need for any post-linkage deprotection steps. A further advantage of the glycosylation method of the

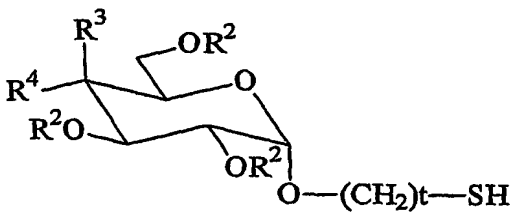
invention is that it allows for the linkage of unprotected carbohydrate moieties to an amino acid, peptide or protein.

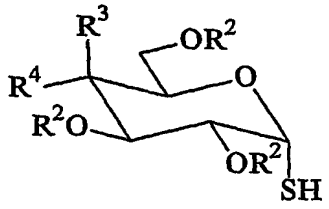
Particularly preferred carbohydrate moieties include $Glc(Ac)_4\beta$-, $Glc(Bn)_4\beta$-, $Gal(Ac)_4\beta$-, $Gal(Bn)_4\beta$-, $Glc(Ac)_4\alpha(1,4)Glc(Ac)_3\alpha(1,4)Glc(Ac)_4\beta$-, β-Glc, β-Gal, α-Glc, α-Gal, β-GlcNAc, β-GalNAc, α-GlcNAc, α-GalNAc, $(Glc\alpha)_6Glc\beta$-, $(Glc\alpha)_2Glc\beta$-, -Et-β-Gal, -Et-β-Glc, Et-α-Glc, -Et-α-Man, -Et-Lac, $-\beta-Glc(Ac)_2$, $-\beta-Glc(Ac)_3$, $-Et-\alpha-Glc(Ac)_2$, $-Et-\alpha-Glc(Ac)_3$, $-Et-\alpha-Glc(Ac)_4$, $-Et-\beta-Glc(Ac)_2$, $-Et-\beta-Glc(Ac)_3$, $-Et-\beta-Glc(Ac)_4$, $-Et-\alpha-Man(Ac)_3$, $-Et-\alpha-Man(Ac)_4$, $-Et-\beta-Gal(Ac)_3$, $-Et-\beta-Gal(Ac)_4$, $-Et-Lac(Ac)_5$, $-Et-Lac(Ac)_6$, $-Et-Lac(Ac)_7$, and their deprotected equivalents.

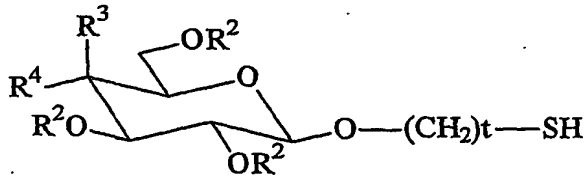
Preferably, any saccharide units making up the carbohydrate moiety which are derived from naturally occurring sugars will each be in the naturally occurring enantiomeric form, which may be either the D-form (e.g. D-glucose or D-galactose), or the L-form (e.g. L-rhamnose or L-fucose). Any anomeric linkages may be α- or β- linkages.

Examples of suitable thio sugars based on galactose and glucose are shown generically below:









wherein:

each R^2 independently denotes H, a saccharide moiety or a suitable protecting group, for example Ac or Bn, preferably each R^2 denotes H;
one of R^3 and R^4 denotes H and the other denotes OH, O-protecting group or O-saccharide moiety, preferably H or O-saccharide moiety; and
t denotes 2 to 10, preferably 2 to 6, more preferably 2 or 3.

Organic compounds containing at least one thiol group, including some thio-sugars, are commercially available or may be prepared using methods known in the art, for example methods for the preparation of thiol compounds in general, and thio-sugars in particular. For example, thio sugars may be prepared from the corresponding halo sugars by treatment of the halo sugar with thiourea to afford the corresponding isothiouronium salt (W. A. Bonner, J. E. Kahn, *J. Am. Chem. Soc.* **1951**, 73) followed by mild hydrolysis with sodium metabisulfite to give the corresponding thiol. If necessary, suitable protecting groups may be used during the synthesis of any thio-sugars.

The reaction of the selenenylsulfide group with the organic compound containing a thiol group (i.e. the disulfide bond forming reaction) is generally carried out in the presence of a buffer at neutral or basic pH (e.g. about pH 7 to about pH 9.5), with slightly basic pHs being preferred (e.g. about pH 8 to about pH 9). Suitable buffers include HEPES, CHES, MES and Tris. The pH should be such that little or no unwanted denaturation of the protein or peptide occurs during the reaction. Similarly, the reaction temperature should be selected to avoid any significant damage to any temperature sensitive compounds. For example, a reaction with a protein or peptide is preferably carried out at ambient temperature or below to avoid any denaturation. Aqueous or organic solvent systems may be used, with aqueous solvent systems being preferred to ensure the dissolution of the protein, amino acid or

peptide. Aqueous solvent systems are also preferred as they allow the use of unprotected carbohydrate compounds as the organic compound. The reaction is generally fairly quick, for example often taking less than 1 hour.

In general, an excess of the organic compound containing at least one thiol group will be used, for example 10-20 equivalents based on the protein, amino acid or peptide. However, as little as 1 mol equivalent may be used in some cases. Carbohydrate compounds may be expensive and time-consuming to obtain in large quantities. Therefore, when the organic compound containing at least one thiol group is a carbohydrate compound, for reasons of economy it is desirable to use the minimum possible number of equivalents. Prior art methods for protein glycosylation often require use of a very large excess of the carbohydrate compound, for example often of the order of 1000 equivalents (B. G. Davis, *Curr. Opin. Biotechnol.* **2003**, *14*, 379). The method of the invention therefore advantageously allows use of fewer equivalents of the glycosyl compound than the prior art methods.

The invention will be further illustrated by the following non-limiting Examples.

General Experimental

Melting points were recorded on a Kofler hot block and are uncorrected. Proton nuclear magnetic resonance (δ_H) spectra 400 MHz spectra were assigned using COSY. Carbon nuclear magnetic resonance (δ_C) spectra were assigned using HMQC. Multiplicities were assigned using DEPT sequence. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard.

Infrared spectra adsorption maxima were recorded in wavenumbers (cm^{-1}) and classified as s (strong) and br (broad). Low resolution mass spectra were recorded using electrospray ionisation (ESI), or using chemical

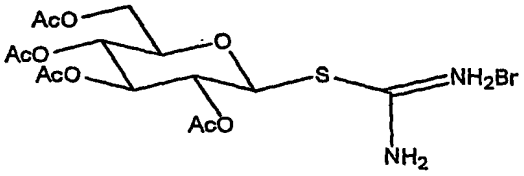
ionization (NH_3, CI) techniques as stated. High resolution mass spectra were recorded using chemical ionization (NH_3, CI) techniques, or using electrospray ionization (NH_3, CI) techniques, or using field ionisation (FI+) as stated. M/z values are reported in Daltons and are followed by their percentage abundance in parentheses.

Optical rotations were measured on a polarimeter with a path length of 1 dm. Concentrations are given in g/100 mL.

Thin layer chromatography (t.l.c) was carried out on Merck Kieselgel $60F_{254}$ pre-coated glassbacked plates. Visulation of the plates was achieved using a u.v lamp (λ_{max} = 254 or 365 nm), and/or ammonium molybdate (5% in 2M H_2SO_4) or sulfuric acid (5% in EtOH). Flash column chromatography was carried out using Sorbsil C60 40/60 silica. Dichloromethane (DCM) was distilled from calcium hydride. Acetone was distilled from anhydrous calcium sulfate. Remaining anhydrous solvents were purchased from Fluka. 'Petrol' refers to the fraction of petroleum ether boiling in the range 40-60°C.

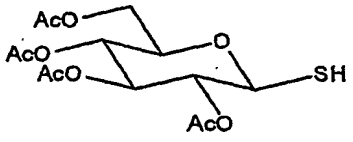
Protein Mass spectrometry: Liquid chromatography/mass spectrometry was performed on a Micromass LCT (ESI-TOF-MS) coupled to a Waters Alliance 2790 HPLC using a Phenomenex Jupiter C5 column (150 x 2.1 mm x 5 μm). Water (solvent A) and acetonitrile (solvent B), each containing 0.5% formic acid, were used as the mobile phase at a flow rate of 0.2 ml min^{-1}. The gradient was programmed as follows: 95% A (3 min isocratic) to 100 % B after 16 min then isocratic for 2 min. The electrospray source of the LCT was operated with a capillary voltage of 3 kV and a cone voltage of 30 V. Nitrogen was used as the nebuliser and desolvation gas at a total flow of 400 l hr^{-1}. Myoglobin (horse heart) was used as a calibration standard and to test the sensitivity of the system.

Example 1 (2,3,4,6-Tetra-O-acetyl-β-D-glucopyranosyl)-1-isothiouronium bromide



2,3,4,6-Tetra-*O*-acetyl-α-D-glucopyranosyl bromide (11.0 g, 26.4 mmol) and thiourea (3.10 g, 41.9 mmol) were dissolved in anhydrous acetone (30 mL) under argon and heated to 60°C. After 20 min a white solid precipitated. The precipitate was removed by filtration, the filtrate was returned to reflux, this process was repeated until the solid ceased to precipitate. The off-white crystals were combined and recrystallised from acetone/petrol to afford the title compound (11.4 g, 76%) as a white crystalline solid mp 194-196°C [Lit. 191°C (H. Beyer, U. Schultz, *Chem. Ber.* 1954, *87*, 78)]; $[\alpha]_D^{25}$ -5.6 (c, 1.0 in H_2O) [Lit. $[\alpha]_D^{25}$ -7.6 (c, 1.4 in H_2O) (W. A. Bonner, J. E. Kahn, *J Am Chem Soc* 1951, *73*, 2241)]; δ_H (400 MHz, DMSO-d_6) 1.97, 2.00, 2.02, 2.06 (12H, 4 x s, 4 x CH_3), 4.06-4.25 (3H, m, H-5, H-6, H-6'), 5.07-5.12 (2H, m, H-2, H-4), 5.31 (1H, at, *J* 9.5 Hz, H-3), 5.77 (1H, d, $J_{1,2}$ 9.9 Hz, H-1), 9.13 (2H, brs, NH_2), 9.29 (2H, brs, NH_2).

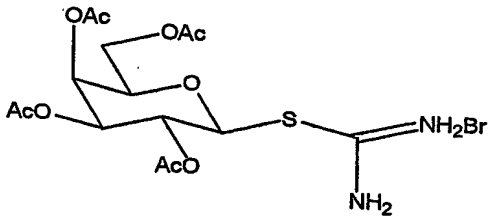
Example 2: 1-Thio-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranose



(2,3,4,6-Tetra-*O*-acetyl-β-D-glucopyranosyl)-1-isothiouronium bromide (9.0 g, 18.8 mmol) and $Na_2S_2O_5$ (4.93 g, 26.0 mmol) were added to a stirred mixture of DCM (150 mL) and water (70 mL). The mixture was heated to reflux under argon. After 1.5 h the reaction was cooled to room temperature (RT) and the phases were

separated. The aqueous layer was re-extracted with DCM (3 x 50 mL). The combined organic layers were washed with water (50 mL), dried over $MgSO_4$, filtered and the solvent removed *in vacuo* to afford the title compound (6.14 g, 90%) as a white solid, mp 112-114°C [Lit. 113-114°C (R. J. Ferrier, R. H. Furneaux, *Carbohydr. Res.* 1977, *57*, 73)]; $[\alpha]_D^{24}$ +6.3 (c, 1.2 in $CHCl_3$) [Lit. $[\alpha]_D^{20}$ +5.0 (c, 1.1 in $CHCl_3$) (R. J. Ferrier, R. H. Furneaux, *Carbohydr. Res.* 1977, *57*, 73)]; δ_H (400 MHz, $CDCl_3$) 1.99, 2.00, 2.05, 2.06 (12H, 4 x s, 4 x CH_3), 2.30 (1H, d, $J_{1,SH}$ 10.2 Hz, SH), 3.71 (1H, ddd, $J_{4,5}$ 10.0 Hz, $J_{5,6}$ 2.4 Hz, $J_{5,6'}$ 4.7 Hz, H-5), 4.10 (1H, dd, $J_{6,6'}$ 12.3 Hz, H-6), 4.22 (1H, dd, H-6'), 4.53 (1H, at, *J* 9.9 Hz, H-1), 4.95 (1H, at, *J* 9.5 Hz, H-2), 5.08 (1H, at, *J* 9.8 Hz, H-4), 5.17 (1H, at, *J* 9.4 Hz, H-3).

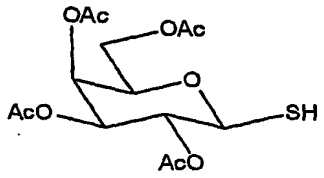
Example 3: (2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl)-1-isothiouronium bromide



2,3,4,6-Tetra-*O*-acetyl-D-β-galactopyranosyl bromide (5.4 g, 13.0 mmol) and thiourea (1.25 g, 16.8 mmol) were dissolved in anhydrous acetone (40 mL) under argon and heated to 60°C. After 1 h the reaction was allowed to cool to room temperature and the resulting residue was filtered and recrystallised from acetone/petrol to afford the title compound (4.6 g, 70%, 2 steps) as a white crystalline solid mp 134-137°C [Lit. 170°C from isopropanol (W. A. Bonner, J. E. Kahn, *J Am Chem Soc* 1951, *73*, 2241)]; $[\alpha]_D^{25}$ +40.4 (c, 1.0 in H_2O) [Lit. $[\alpha]_D^{25}$

+16.0 (c, 1.6 in EtOH, (W. A. Bonner, J. E. Kahn, *J Am Chem Soc* 1951, 73, 2241)); δ_H (500 MHz, DMSO-d_6) 1.96, 2.02, 2.09, 2.15 (12H, 4 x s, 4 x CH_3) 4.06-4.13 (2H, m, H-6, H-6'), 4.45 (1H, t, *J* 6.2 Hz, H-5), 5.12 (1H, at, *J* 9.9 Hz, H-2), 5.24 (1H, dd, $J_{2,3}$ 10.0 Hz, $J_{3,4}$ 3.6 Hz, H-3), 5.39 (1H, d, $J_{3,4}$ 3.1 Hz, H-4), 5.71 (1H, d, $J_{1,2}$ 10.2 Hz, H-1), 9.12, 9.36 (2 x 2H, 2 x brs, 2 x NH_2).

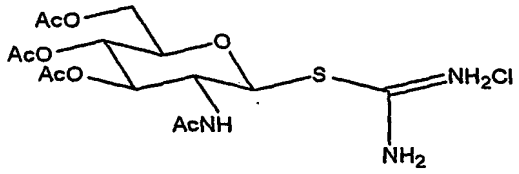
Example 4: 1-Thio-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranose



(2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl)-1-isothiouronium bromide (4.4 g, 8.8 mmol) and $Na_2S_2O_5$ (2.02 g, 10.6 mmol) were added to a stirred mixture of DCM (60 mL) and water (30 mL). The mixture was heated to reflux under argon. After 2.5 h the reaction was cooled to RT and the phases were separated. The aqueous layer was re-extracted with DCM (3 x 50 mL). The combined organic layers were washed with water (100 mL), brine (100 mL), dried over $MgSO_4$, filtered and the solvent removed *in vacuo* to afford the title compound (2.65 g, 81%) as a white solid, mp 83-84°C [Lit. 86.5-88°C (J. Frgala, M. Cerny, J. Stanek, *Collect. Czech. Chem. Commun.* 1975, *40*, 1411)]; $[\alpha]_D^{24}$ +30.1 (c, 1.0 in $CHCl_3$) [Lit. $[\alpha]_D^{19}$ +32.0 (c, 3.5 in $CHCl_3$) (J. Frgala, M. Cerny, J. Stanek, *Collect. Czech. Chem. Commun.* 1975, *40*, 1411)]; δ_H (400 MHz, $CDCl_3$) 1.99, 2.06, 2.10, 2.17 (12H, 4 x s, 4 x CH_3), 2.38 (1H, d, $J_{1,SH}$ 10.3 Hz, S<u>H</u>), 3.95 (1H, dt, $J_{4,5}$ 1.2 Hz, $J_{5,6}$ 6.6 Hz, $J_{5,6'}$ 6.6 Hz, H-5), 4.09-4.14 (2H, m, H-6, H-6'), 4.53 (1H, at, *J* 9.9 Hz,

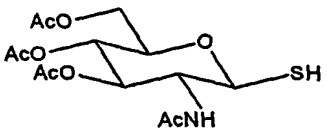
H-1), 5.02 (1H, dd, $J_{2,3}$ 10.1, $J_{3,4}$ 3.4 Hz, H-3), 5.19 (1H, at, J 10.0 Hz, H-2), 5.44 (1H, at, dd, $J_{3,4}$ 3.7 Hz, $J_{4,5}$ 1.2 Hz, H-4).

Example 5: (3,4,6-Tri-*O*-acetyl-2-acetamido-2-deoxy-β-D-glucopyranosyl)-1-isothiouronium chloride



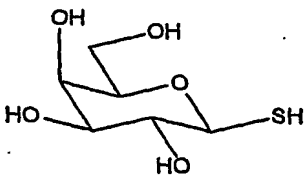
3,4,6-Tri-*O*-acetyl-2-acetamido-2-deoxy-α-D-glucopyranosoyl chloride (3.0 g, 8.2 mmol) and thiourea (1.21 g, 14.6 mmol) were dissolved in anhydrous acetone (25 mL) under argon and heated to 60°C. After 2 h a white solid precipitated. The precipitate was removed by filtration, the filtrate was returned to reflux, this process was repeated until the solid ceased to precipitate. The off white crystals were combined and recrystallised from acetone/petrol to afford (the title compound (2.19 g, 61%) as a white crystalline solid mp 134-137°C [Lit. 179-18°C from EtOH (D. Horton, M. L. Wolfrom, *Journal of Organic Chemistry* 1962, *27*, 1794)]; $[\alpha]_D^{25}$ -25.2 (c, 1.0 in H_2O) [Lit. $[\alpha]_D^{25}$ -29.3 (c, 1.1 in MeOH) (D. Horton, M. L. Wolfrom, *Journal of Organic Chemistry* 1962, *27*, 1794)]; δ_H (400 MHz, DMSO-d_6) 1.80 (3H, s, NHCOC$\underline{H}_3$), 1.94, 1.98, 2.08 (9H, 3 x s, 3 x CH_3), 4.05 (1H, dd, $J_{5,6}$ 2.4 Hz, $J_{6,6'}$ 12.4 Hz, H-6), 4.17 (1H, dd, $J_{5,6'}$ 5.0 Hz, $J_{6,6'}$ 12.3 Hz, H-6'), 4.26 (1H, ddd, $J_{4,5}$ 10.2 Hz, $J_{5,6}$ 2.3 Hz, $J_{5,6'}$ 4.7 Hz, H-5), 4.93 (1H, at, J 9.9 Hz, H-4), 5.12 (1H, at, J 9.9 Hz, H-3), 5.73 (1H, d, $J_{1,2}$ 10.4 Hz, H-1), 8.48 (1H, d, J 4.7 Hz, N$\underline{H}$Ac), 9.13 (2H, brs, NH_2), 9.29 (2H, brs, NH_2).

Example 6: 1-Thio-3,4,6-tri-O-acetyl-2-acetamido-2-deoxy-β-D-glucopyranose



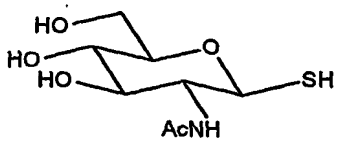
(3,4,6-Tri-O-acetyl-2-acetamido-2-deoxy-β-D-glucopyranosyl)-1-isothiouronium chloride (1.75 g, 39.8 mmol) and $Na_2S_2O_5$ (0.91 g, 4.8 mmol) were added to a stirred mixture of DCM (30 mL) and water (15 mL). The mixture was heated to reflux under argon. After 2 h the reaction was cooled to RT and the phases were separated. The aqueous layer was re-extracted with DCM (2 x 50 mL). The combined organic layers were washed with water (50 mL), brine (50 mL), dried over $MgSO_4$, filtered and the solvent removed *in vacuo*. Recrystalisation from EtOAc/petrol afforded the title compound (1.00 g, 68%) as a white solid, mp 165-167°C [Lit. 167-168°C (W. M. zu Reckendorf, W. A. Bonner, *Journal of Organic Chemistry* 1961, *26*, 4596)]; $[\alpha]_D^{25}$ -24.8 (c, 1.0 in $CHCl_3$) [Lit. $[\alpha]_D^{25}$ -14.5 (c, 0.9 in $CHCl_3$) (W. M. zu Reckendorf, W. A. Bonner, *Journal of Organic Chemistry* 1961, *26*, 4596)]; δ_H (400 MHz, $CDCl_3$) 1.99, 2.03, 2.05, 2.10 (12H, 4 x s, 4 x CH_3), 2.57 (1H, d, $J_{1,SH}$ 9.2 Hz, SH), 3.67 (1H, ddd, $J_{4,5}$ 9.7 Hz, $J_{5,6}$ 4.8 Hz, $J_{5,6'}$ 2.3 Hz, H-5), 4.09-4.17 (2H, m, H-2, H-3), 4.24 (1H, dd, $J_{5,6}$ 4.8 Hz, $J_{6,6'}$ 12.4 Hz, H-6), 4.59 (1H, at, *J* 9.8 Hz, H-1), 5.06-5.15 (2H, m, H-4, H-6'), 5.72 (1H, d, *J* 9.2 Hz, NH).

Example 7: 1-Thio-β-D-galactopyranose



1-Thio-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranose (3.00 g, 7.3 mmol) and NaOMe (40 mg, 0.73 mmol) were added to a stirred solution of MeOH (40 ml). After 2 h, t.l.c. (EtOAc/petrol 1:1) indicated the formation of a product (R_f 0.0) with complete consumption of the starting material (R_f 0.5). The reaction was neutralised with the addition of Dowex®-50 ion exchange resin after which point the reaction was filtered and concentrated *in vacuo*. Recrystalisation from MeOH/EtOAc afforded the title compound (1.41 g, 98%) as a white crystalline solid m.p. 100-102°C; $[\alpha]_D^{22}$ +47.6 (c, 1.0 in MeOH; δ_H (400 MHz, CD_3OD), 2.62 (1H, d, $J_{1,SH}$ 8.3 Hz, SH), 3.47-3.49 (2H, m, H-2, H-3), 3.57 (1H, at, J 5.9 Hz, H-5), 3.68 (1H, dd, $J_{5,6}$ 5.0 Hz, $J_{6,6'}$ 11.4 Hz, H-6), 3.75 (1H, dd, $J_{5,6'}$ 6.9 Hz, $J_{6,6'}$ 11.5 Hz, H-6'), 3.91 (1H, bs, H-4), 4.37 (1H, bd, J 7.7 Hz, H-1); δ_C (100 MHz, CD_3OD), 61.6 (t, C-6), 69.6 (d, C-4), 74.4, 74.8 (2 x d, C-2, C-3), 80.1 (d, C-5), 81.4 (d, C-1); m/z (ES-) 196 (100%, $M-H^+$); m/z HRMS (ES-) Calcd. for $C_6H_{12}O_5S$ ($M-H^+$) 195.0327. Found 195.0323.

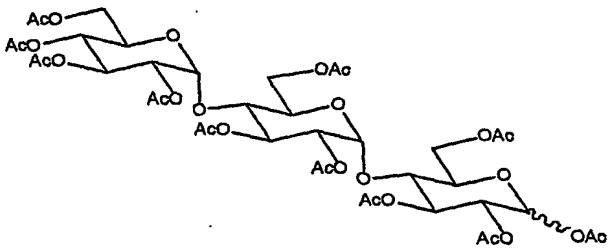
Example 8: 1-Thio-2-acetamido-2-deoxy-β-D-glucopyranose



3,4,6-Tri-*O*-acetyl-2-acetylamino-2-deoxy-β-D-glucopyranosyl thiol (400 mg, 0.98 mmol) and sodium methoxide (18 mg, 0.03 mmol) were added to a stirred solution of methanol (5ml). After a 30 min period, t.l.c. (ethyl acetate) indicated the formation of a product (R_f 0.0) with complete consumption of the starting material (R_f 0.2). The reaction was neutralised with the addition of Dowex®-50 ion exchange resin after

which point the reaction was filtered and concentrated *in vacuo*. Recrystallisation from methanol/ethyl acetate afforded the title product (13.35 g, 95%) as a white crystalline solid; m.p. 85-88°C [Lit. 86-88°C][18]; $[\alpha]_D^{22}$ -10.4 (c, 1.0 in MeOH) [Lit. $[\alpha]_D^{25}$ +177.1 (c, 1.45 in $CHCl_3$)][18]; δ_H (400 MHz, MeOH), 2.00 (3H, s, CH_3), 3.27-3.37 (2H, m, H-4, H-5), 3.42 (1H, at *J* 9.1 Hz, H-3), 3.64-3.73 (2H, m, H2, H-6), 3.87 (1H, dd, $J_{5,6}$ 2.1 Hz, $J_{6,6'}$ 12.0 Hz, H-6'), 4.56 (1H, d, $J_{1,2}$ 10.0 Hz, H-1), 8.11 (1H, bd, $J_{NH,2}$ 9.1 Hz, NH).

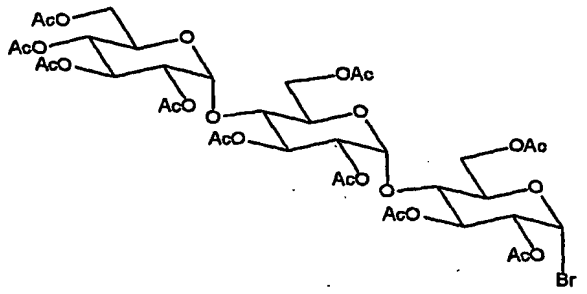
Example 9: 1,2,3,6-tetra-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*D*-glucopyranosyl)-α-D-glucopyranosyl)-D-glucopyranose



Sodium acetate (700 mg, 8.3 mmol) was added to acetic anhydride (50 mL) and heated to reflux, at which point maltotriose (3.00 g, 6.0 mmol) was added and stirred vigorously. After 90 min, t.l.c. (petrol:ethyl acetate, 1:2) indicated the formation of a product (R_f 0.3) with complete consumption of the starting material (R_f 0.0). The reaction was allowed to cool to RT and diluted with DCM (50 mL) and partitioned with water (100 mL). The phases were separated and the aqueous layer was re-extracted with DCM (2 x 50 mL). The combined organic layers were washed with sodium hydrogen carbonate (400 mL of a saturated aqueous solution) until pH 8 was obtained, brine (200 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford the title product as a mixture of anomers (α/β, 2/11) as an amorphous white

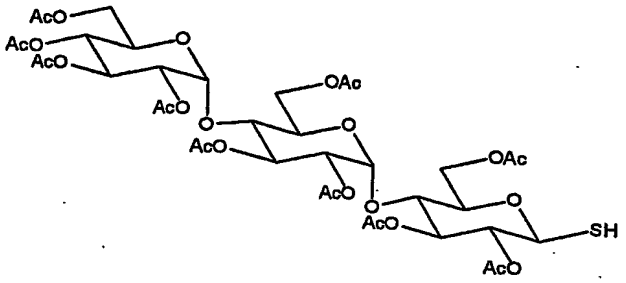
solid; for β compound: δ_H (500 MHz, $CDCl_3$) 2.05, 2.07, 2.10, 2.14, 2.15, 2.19, 2.21, 2.27 (30H, 8 x s, 10 x OAc), 3.92 (1H, ddd, $J_{4,5}$ 9.5 Hz, $J_{5,6}$ 2.9 Hz, $J_{6,6}$ 4.1 Hz, H-5a), 3.95-4.01 (3H, m, H-4b, H-5b, H-5c), 4.05 (1H, at, J 9.1 Hz, H-4a), 4.09 (1H, dd, $J_{5,6}$ 2.5 Hz, $J_{6,6'}$ 12.7 Hz, H-6c), 4.21 (1H, dd, $J_{5,6}$ 3.4 Hz, $J_{6,6'}$ 12.6 Hz, H-6b), 4.29 (1H, dd, $J_{5,6}$ 3.4 Hz, $J_{6,6'}$ 12.4 Hz, H-6'c), 4.35 (1H, dd, $J_{5,6}$ 4.3 Hz, $J_{6,6'}$ 12.3 Hz, H-6a), 4.48-4.52 (2H, m, H-6'a, H-6'b), 4.78 (1H, dd, $J_{1,2}$ 4.1 Hz, $J_{2,3}$ 10.3 Hz, H-2b), 4.90 (1H, dd, $J_{1,2}$ 4.1 Hz, $J_{2,3}$ 10.6 Hz, H-2c), 5.01 (1H, dd, $J_{1,2}$ 8.0 Hz, $J_{2,3}$ 9.0 Hz, H-2a), 5.11 (1H, at, J 10.1 Hz, H-4c), 5.31 (1H, d, $J_{1,2}$ 3.9 Hz, H-1b), 5.32-5.44 (3H, m, H-3a, H-3b, H-3c), 5.45 (1H, d, $J_{1,2}$ 4.1 Hz, H-1c), 5.79 (1H, d, $J_{1,2}$ 8.2 Hz, H-1a); for α compound selected data only: δ_H (500 MHz, $CDCl_3$) 2.08, 2.09, 2.12, 2.18, 2.21, 2.23, 2.26 (30H, 8 x s, 10 x OAc), 5.07 (1H, at, J 9.9 Hz), 6.28 (1H, d, $J_{1,2}$ 3.8 Hz, H-1a). Remaining signals lie in the following multiplet regions, 3.85-3.89, 3.90-3.98, 3.99-4.07, 4.15-4.18, 4.23-4.27, 4.29-4.32, 4.43-4.49, 4.74-4.76, 4.84-4.87, 4.98-4.94, 5.25-5.54; m/z (ES+) 984 (MNH_4^+, 30%), 989 (MNa^+, 100%); m/z HRMS (ES^+) Calcd. For $C_{40}H_{58}O_{27}N$ (MNH_4^+) 984.3196 Found 984.3199.

Example 10: 2,3,6-Tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*D*-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl bromide



1,2,3,6-Tetra-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-D-glucopyranose (200 mg, 0.21 mmol) was dissolved in anhydrous DCM (5 mL). To this hydrogen bromide (33% in acetic acid, 2 mL) was added. The mixture was left under argon at RT. After a 30 min period, t.l.c. (petrol:ethyl acetate, 1:2) indicated the formation of a product (R_f 0.6) with complete consumption of the starting material (R_f 0.3). The reaction mixture was partitioned between DCM (10 mL) and water (10 mL), and the aqueous layer re-extracted with DCM (3 x 10 mL). The combined organic layers were washed with sodium hydrogen carbonate (20 mL of a saturated aqueous solution) until pH 8 was obtained, brine (20 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford the title product (203 mg, 98%) as a white foam; $[\alpha]_D^{22}$ +152.2 (c, 1.0 in $CHCl_3$); δ_H (400 MHz, $CDCl_3$) 2.03, 2.05, 2.06, 2.08, 2.10, 2.13, 2.18, 2.21 (30H, 10 x $COCH_3$), 3.93-3.99 (3H, m, H-4b, H-5a, H-5b), 4.05-4.10 (2H, m, H-4c, H-6a), 4.20 (1H, dd, $J_{5,6}$ 1.8 Hz, $J_{6,6'}$ 12.2 Hz, H-6b), 4.26-4.34 (2H, m, H-5c, H-6a'), 4.35 (1H, dd, $J_{5,6}$ 3.5 Hz, $J_{6,6'}$ 12.7 Hz, H-6c), 4.52 (1H, dd, $J_{5,6}$ 0.6 Hz, $J_{6,6'}$ 12.2 Hz, H-6b'), 4.57 (1H, dd, $J_{5,6}$ 2.1 Hz, $J_{6,6'}$ 12.4 Hz, H-6c''), 4.74 (1H, dd, $J_{1,2}$ 4.1 Hz, $J_{2,3}$ 9.9 Hz, H-2c), 4.78 (1H, dd, $J_{1,2}$ 4.2 Hz, $J_{2,3}$ 10.2 Hz, H-2b), 4.88 (1H, dd, $J_{1,2}$ 4.0 Hz, $J_{2,3}$ 10.5 Hz, H-2a), 5.10 (1H, at, *J* 9.7 Hz, H-4a), 5.32 (1H, d, $J_{1,2}$ 4.0 Hz, H-1b), 5.39 (1H, at, *J* 9.9 Hz, H-3q), 5.43-5.46 (1H, m, H-3b), 5,45 (1H, d, $J_{1,2}$ 3.8 Hz, H-1a), 5.64 (1H, at, *J* 9.5 Hz, H-3c), 6.53 (1H, d, $J_{1,2}$ 3.9 Hz, H-1c).

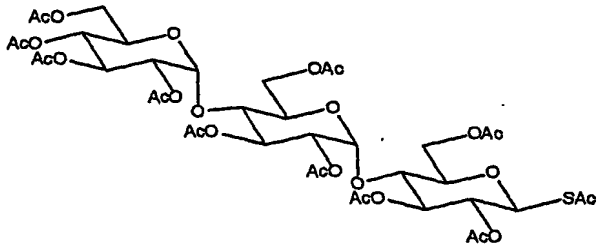
Example 11: 1-Thio-2,3,6-Tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-β-D-glucopyranose



2,3,6-Tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl bromide (2.10 g, 2.10 mmol) was dissolved in anhydrous acetone (60 mL). To this anhydrous thiourea (315 mg, 4.2 mmol) was added and then heated to reflux under an atmosphere of argon. After a 6.5 h period, t.l.c. (petrol:ethyl acetate, 1:2) indicated the formation of a product (R_f 0.0) with complete consumption of the starting material (R_f 0.3). The reaction was concentrated *in vacuo* and titurated with DCM to remove the organics from the excess thiourea. The filtrate was concentrated *in vacuo* and the residue was purified by column flash chromatography (ethyl acetate/methanol, 9:1) to afford the intermediate 2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-β-D-glucopyranosyl-1-isothiouronium bromide (1.14g, 50%) which was carried on without characterisation. This intermediate (100 mg, 0.09 mmol) and $Na_2S_2O_5$ (22 mg, 0.11 mmol) were added to a stirred mixture of DCM (30 mL) and water (15 mL). The mixture was heated to reflux under argon. After 2.5 h, t.l.c. (petrol:ethyl acetate, 1:2) indicated the formation of a product (R_f 0.4) with complete consumption of the starting material (R_f 0.0), at which point the reaction was cooled to RT and the phases separated. The aqueous layer was re-extracted with DCM (2 x 20 mL). The combined organic layers were washed with brine (20 mL), dried ($MgSO_4$), filtered and the solvent removed *in vacuo* to afford the

title product (74 mg, 84%) as a white amorphous solid; $[\alpha]_D^{22}$ +99.5 (c, 1.0 in $CHCl_3$); δ_H (400 MHz, $CDCl_3$) 1.99, 2.00, 2.01, 2.02, 2.03, 2.05, 2.10, 2.15, 2.18 (30H, 9 x s,10 x $COCH_3$), 3.72-3.76 (1H, m, H-5a), 3.90-4.00 (4H, m, H-4a, H-4b, H-5b, H-5c), 4.05 (1H, dd, $J_{5,6}$ 2.2 Hz, $J_{6,6'}$ 12.3 Hz, H-6c), 4.17 (1H, dd, $J_{5,6}$ 3.3 Hz, $J_{6,6'}$ 12.3 Hz, H-6b), 4.25 (1H, dd, $J_{5,6}$ 3.6 Hz, $J_{6,6'}$ 12.5 Hz, H-6c'), 4.30 (1H, $J_{5,6}$ 4.3 Hz, $J_{6,6'}$ 12.2 Hz, H-6c), 4.44 (1H, dd, $J_{5,6}$ 2.2 Hz, $J_{6,6'}$ 12.1 Hz, H-6a'), 4.46 (1H, dd, $J_{5,6}$ 2.2 Hz, $J_{6,6'}$12.2 Hz, H-6b'), 4.59 (1H, d, $J_{1,2}$ 9.7 Hz, H-1a), 4.74 (1H, dd, $J_{1,2}$ 4.1 Hz, $J_{2,3}$ 10.6 Hz, H-2b), 4.80 (1H, at, J 9.0 Hz, H-2a), 4.85 (1H, dd, $J_{1,2}$ 4.1 Hz, $J_{2,3}$ 10.6 Hz, H-2c), 5.07 (1H, at, J 9.9 Hz, H-4c), 5.25 (1H, at, J 9.0 Hz, H-3a), 5.26 (1H, d, $J_{1,2}$ 4.1 Hz, H-1b), 5.35 (1H, at, J 10.0 Hz, H-3b), 5.37-5.41 (2H, m, H-1c, H-3c).

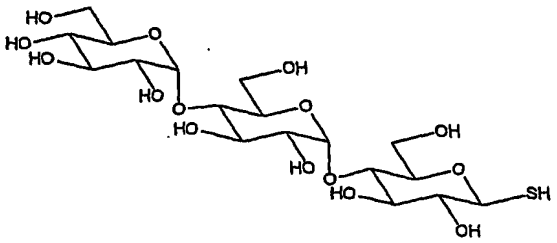
Example 12: 1-Thioacetyl-2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-β-D-glucopyranose



2,3,6-Tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)- α-D-glucopyranosyl)-β -D-glucopyranosyl bromide (11.2 g, 11.6 mmol) and potassium thioacetate (3.96 g, 34.8 mmol) were suspended in anhydrous THF (40 ml) and heated to reflux under an inert atmosphere of argon. After 14 h, t.l.c. (petrol/EtOAc, 1:2) indicated the formation of a major

product (R_f 0.4) along with complete consumption of the starting material (R_f 0.45). The reaction was diluted with water (80 mL) and allowed to cool to RT. The phases were separated and the aqueous phase was re-extracted with DCM (3 x 40 mL). The combined organic layers were washed with sat. $NaHCO_3$ (50 mL) until pH 8 was obtained, brine (50 mL), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (petrol/EtOAc, 1:4) to afford the title compound (8.08 g, 71%) as a white foam; $[\alpha]_D^{25}$ +86.4 (c, 1.0 in $CHCl_3$); δ_H (400 MHz, $CDCl_3$) 2.01, 2.02, 2.05, 2.08, 2.11, 2.17 (27H, 6 x s, 9 x OAc), 2.40 (3H, s, SAc), 3.88 (1H, ddd, $J_{4,5}$ 9.8 Hz, $J_{5,6}$ 4.0 Hz, $J_{5,6'}$ 2.7 Hz, H-5a), 3.92-4.01 (4H, m, H-4a, H-4b, H-5b, H-5c), 4.07 (1H, dd, $J_{5,6}$ 2.4 Hz, $J_{6,6'}$ 12.3 Hz, H-6c), 4.19 (1H, dd, $J_{5,6}$ 3.5 Hz, $J_{6,6'}$ 12.2 Hz, H-6b), 4.27 (1H, dd, $J_{5,6'}$ 3.8 Hz, $J_{6,6'}$ 12.3 Hz, H-6'c), 4.30 (1H, dd, $J_{5,6}$ 4.2 Hz, $J_{6,6'}$ 12.4 Hz, H-6a), 4.46 (1H, dd, $J_{5,6'}$ 2.6 Hz, $J_{6,6'}$ 12.3 Hz, H-6'b), 4.47 (1H, dd, $J_{5,6'}$ 2.2 Hz, $J_{6,6'}$ 12.2 Hz, H-6'a), 4.76 (1H, dd, $J_{1,2}$ 3.9 Hz, $J_{2,3}$ 10.0 Hz, H-2b), 4.87 (1H, dd, $J_{1,2}$ 3.8 Hz, $J_{2,3}$ 10.6 Hz, H-2c), 5.99 (1H, dd, $J_{1,2}$ 10.3 Hz, $J_{2,3}$ 9.1 Hz, H-2a), 5.08 (1H, at, J 9.9 Hz, H-4c), 5.27 (1H, d, $J_{1,2}$ 4.0 Hz, H-1b), 5.31 (1H, d, $J_{1,2}$ 10.0 Hz, H-1a), 5.33-5.43 (4H, m, H-1c, H-3a, H-3b, H-3c); δ_C (125 MHz, $CDCl_3$) 20.7, 20.8, 20.9, 21.0, 21.1 (5 x q, 10 x $CO\underline{C}H_3$, $SCOCH_3$), 31.0 (q, $SCO\underline{C}H_3$) 61.9 (t, C-6c), 62.7 (t, C-6b), 63.3 (t, C-6a), 68.4 (d, C-4c), 69.0 (d, C-5b), 69.5 (d, C-5c), 69.8 (d, C-3c), 70.3 (d, C-2a), 70.5 (d, C-2c), 70.9 (d, C-2a), 72.1 (d, C-3b), 73.0 (d, C-4b), 74.1 (d, C-4a), 76.6 (d, C-3a), 76.9 (d, C-5a), 80.2 (d, C-1a), 96.1 (d, C-1c), 96.4 (d, C-1b), 169.4, 169.6, 169.8, 169.9, 170.3, 170.5, 170.6 (7 x s, 10 x $\underline{C}OCH_3$), 196.0 (s, $S\underline{C}OCH_3$); m/z (ES+) 1000 (MNH_4^+, 60%), 1003 (MNa^+, 100%).

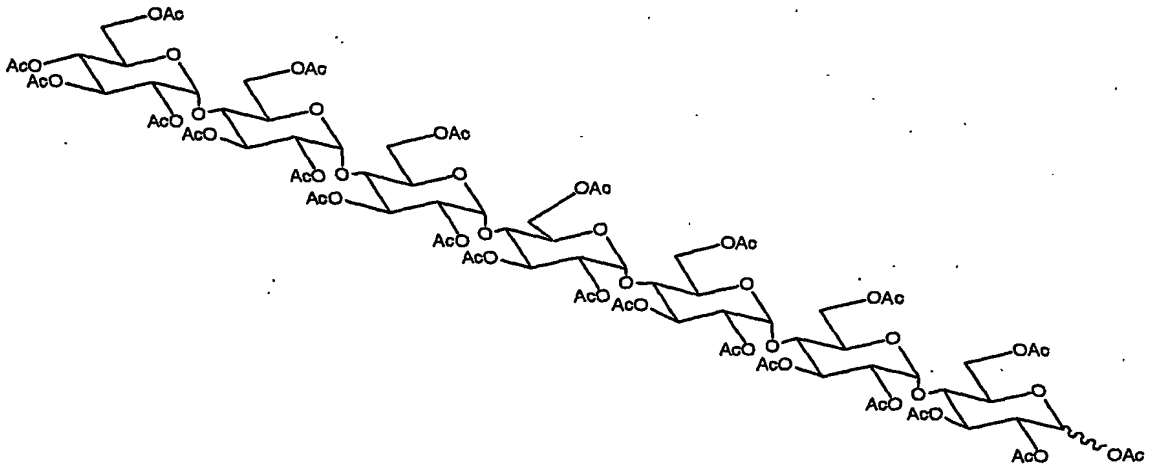
Example 13: 1-Thio-β-D-maltotriose



1-Thioacetyl-2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)- α-D-glucopyranosyl)-1-thio-β-D-glucopyranose (600 mg, 0.6 mmol) and NaOAc (18 mg, 0.18 mmol) were added to a stirred solution of MeOH (10 ml). After 10 min, t.l.c. (EtOAc/MeOH, 9:1) indicated the formation of a product (R_f 0.0) with complete consumption of the starting material (R_f 0.9). The reaction was neutralised with the addition of Dowex®-50 ion exchange resin after which point the reaction was filtered and concentrated *in vacuo* to afford the title compound (305 mg, 98%) as an amorphous solid; $[\alpha]_D^{25}$ +123 (c, 1.0 in MeOH); δ_H (400 MHz, D_2O), 3.15 (1H, at, *J* 9.2 Hz, H-2a), 3.26 (1H, at, *J* 9.3 Hz), 3.41-3.82 (16H, m, H-2b, H-2c, H-3a, H-3b, H-3c, H-4a, H-4b, H-4c, H-5a, H-5b, H-5c, H-6a, H-6b, H-6c, H-6'a, H-6'b, H-6'c), 4.42 (1H, d, $J_{1,2}$ 9.6 Hz, H-1a), 5.23 (1H, d, $J_{1,2}$ 1.7 Hz, H-1), 5.24 (1H, d, $J_{1,2}$ 1.8 Hz, H-1); δ_C (100 MHz, D_2O), 60.8, 70.0 (2 x t, C-6a, C-6b, C-6c), 69.6, 71.5, 71.8, 72.1, 73.0, 73.2, 73.6, 76.0, 77.1, 77.6, 79.0 (11 x d, C-2a, C-2b, C-2c, C-3a, C-3b, C-3c, C-4a, C-4b, C-4c, C-5a, C-5b, C-5c), 80.2 (d, C-1a), 99.8, 100.1 (2 x d, C-1b, C-1c); m/z (ES-) 519 (100%, M-H$^+$); m/z HRMS (ES-) calcd. for $C_{18}H_{31}O_{15}S$ (M-H$^+$) 519.1384. Found 519.1389.

<u>Example 14: 1, 2,3,6-Tetra-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-</u>

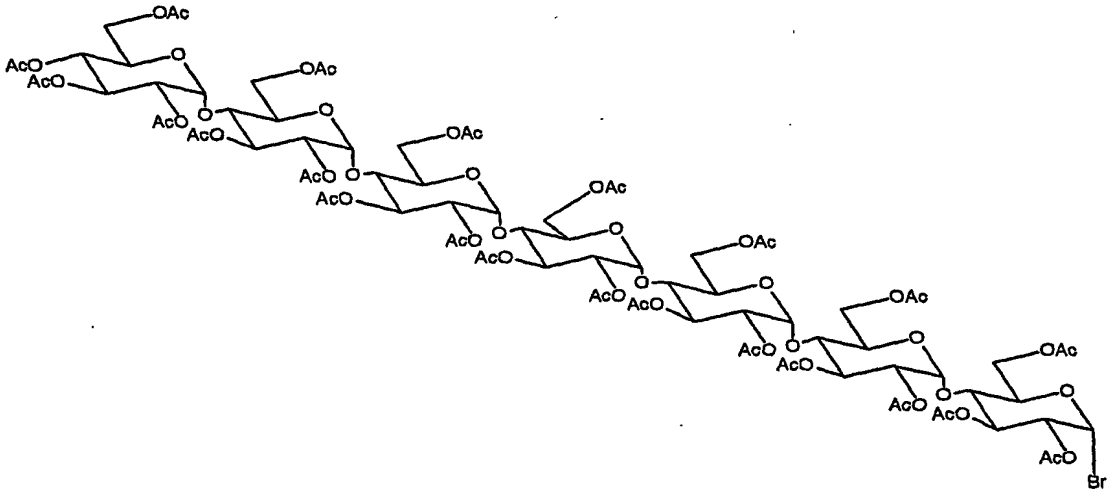
(2,3,4,6-tetra-O-acetyl-α-O-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-D-glucopyranose



Sodium acetate (420 mg, 5.2 mmol) was added to acetic anhydride (30 mL) and heated to reflux, at which point maltoheptose (1.00 g, 0.86 mmol) was added and the reaction stirred vigorously. After 90 min t.l.c. (petrol:ethyl acetate, 1:3) indicated the formation of a product (R_f 0.3) with complete consumption of the starting material (R_f 0.0). The reaction was allowed to cool to RT, diluted with DCM (50 mL) and partitioned with water (100 mL). The phases were separated and the aqueous layer was re-extracted with DCM (2 x 40 mL). The combined organic layers were washed with sodium hydrogen carbonate (200 mL of a saturated aqueous solution) until pH 8 was obtained, brine (100 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (petrol:ethyl acetate, 1:3) to afford the title product as a mixture of anomers as an amorphous white solid (α/β, 15/85); δ_H (500 MHz, $CDCl_3$) 2.02, 2.03, 2.04, 2.05, 2.06, 2.07, 2.08, 2.10, 2.13, 2.19, 2.22, 2.24 (66H, 12 x s, 22 x OAc), 3.89-4.14 (13H, m, H-4a, H-4b, H-4c, H-4d, H-4e, H-4f, H-5a, H-5b, H-5c, H-5d, H-5e, H-5f, H-5g), 4.25-4.34, 4.39 (1H, dd, *J* 4.0 Hz, *J* 12.3 Hz), 4.52-4.56 (13H, m, H-6a, H-6a', H-6b, H-6b', H-6c, H-6c', H-6d, H-6d',

H-6e, H-6e', H-6f, H-6f', H-6d, H-6g'), 4.75-4.79 (5H, m, H-2b, H-2c, H-2d, H-2e, H-2e, H-2f), 4.90 (1H, dd, $J_{1,2}$ 3.7 Hz, $J_{2,3}$ 10.5 Hz, H-2g), 5.00 (1H, at, J 9.4 Hz, H-4g), 5.31-5.45 (13H, m, H-3a, H-3b, H-3c, H-3d, H-3e, H-3f, H-3g, H-1b, H-1c, H-1d, H-1e, H-1f, H-1g), 5.79 (0.85H, d, $J_{1,2}$ 8.1 Hz, H-1aβ), 6.28 (0.15H, d, $J_{1,2}$ 4.0 Hz, H-1aα).

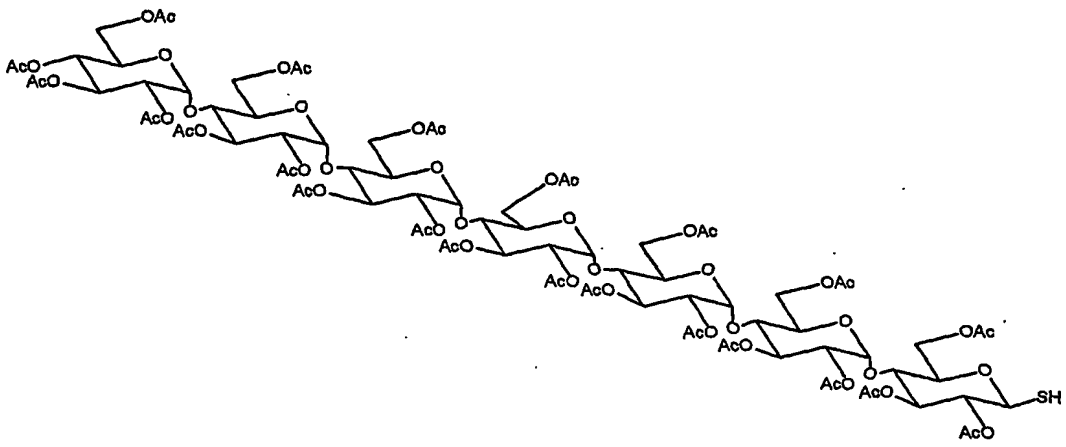
Example 15: 2,3,6-Tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,4,6-tetra-O-acetyl-α-O-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)- α-D-glucopyranosyl)-α-D-glucopyranosyl bromide



1,2,3,6-Tetra-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,4,6-tetra-O-acetyl-α-O-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-D-glucopyranose (100 mg, 0.05 mmol) was dissolved in anhydrous DCM (5 mL). To this hydrogen bromide (33% in acetic acid, 0.5 mL) was added. The mixture was left stirring under an atmosphere of argon

at RT. After a 40 min period, t.l.c. (petrol:ethyl acetate, 1:3) indicated the formation of a product (R_f 0.7) with complete consumption of the starting material (R_f 0.3). The reaction mixture was partitioned between DCM (10 mL) and water (10 mL), and the aqueous layer re-extracted with DCM (3 x 10 mL). The combined organic layers were washed with sodium hydrogen carbonate (150 mL of a saturated aqueous solution) until pH 7 was obtained, brine (20 mL), dried ($MgSO_4$), filtered and concentrated *in vacuo* to afford the title product (98 mg, 96%) as a white foam; $[\alpha]_D^{22}$ +162.0 (c, 1.0 in $CHCl_3$); δ_H (400 MHz, $CDCl_3$) 2.02, 2.03, 2.04, 2.06, 2.08, 2.10, 2.11, 2.14, 2.19, 2.23, 2.24, 2.25 (66H, 12 x s, 22 x OAc), 3.94-4.04 (12H, m, H-4b, H-4c, H-4d, H-4e, H-4f, H-5b, H-5c, H-5d, H-5e, H-5f, H-5g), 4.08 (1H, dd, $J_{5,6}$ 2.2 Hz, $J_{6,6'}$ 12.6 Hz, H-6), 4.19-4.33, 4.53-4.60 (12H, m, H-5a, H-6b, H-6b', H-6c, H-6c', H-6d, H-6d', H-6e, H-6e', H-6f, H-6f', H-6g, H-6g'), 4.12 (1H, at, J 9.5 Hz, H-4a), 4.40 (1H, dd, $J_{5,6}$ 3.1 Hz, $J_{6,6'}$ 12.7 Hz, H-6a), 4.64 (1H, dd, $J_{5,6}$ 2.3 Hz, $J_{6,6'}$ 12.5 Hz, H-6a'), 4.74 (1H, dd, $J_{1,2}$ 3.9 Hz, $J_{2,3}$ 9.7 Hz, H-2a), 4.75-4.97 (5H, m, H-2b, H-2c, H-2d, H-2e, H-2f), 4.89 (1H, d, $J_{1,2}$ 4.0 Hz, $J_{2,3}$ 10.6 Hz, H-2g), 5.11 (1H, at, J 9.9 Hz, H-4g), 5.32-5.47 (12H, m, H-1b, H-1c, H-1d, H-1e, H-1f, H-1g, H-3b, H-3c, H-3d, H-3e, H-3f, H-3g), 5.65 (1H, at, J 9.4 Hz, H-3a), 6.54 (1H, d, $J_{1,2}$ 4.3 Hz, H-1a).

Example 16: 1-Thio-2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,6-tri-O-acetyl-4-O-(2,3,4,6-tetra-O-acetyl-α-O-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)- α-D-glucopyranosyl)-β-D-glucopyranose



2,3,6-Tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl bromide (1.08 g, 0.5 mmol) and tetrabutylammonium iodide (19 mg, 0.05 mmol) were dissolved in anhydrous acetone (50 mL). To this dried thiourea (52 mg, 0.7 mmol) was added and the reaction was then heated to reflux under an atmosphere of argon. After a 8 h period, t.l.c. (petrol:ethyl acetate, 1:4) indicated the formation of a minor product (R_f 0.0) with complete consumption of the starting material (R_f 0.6). The reaction was concentrated *in vacuo* and titurated with DCM to remove the organics from the excess thiourea. The filtrate was concentrated *in vacuo* and the residue was purified by column flash chromatography (ethyl acetate/methanol, 9:1) to afford the intermediate 2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-β-D-glucopyranosyl-1-isothiouronium bromide (212 mg, 19%) which was taken on further without characterisation.

This intermediate (210 mg, 0.09 mmol) and $Na_2S_2O_5$ (22 mg, 0.11 mmol) were added to a stirred mixture of DCM (10 mL) and water (5 mL). The mixture was heated to reflux under argon. After 4.5 h, t.l.c. (petrol:ethyl acetate, 1:2) indicated the formation of a product (R_f 0.2) with complete consumption of the starting material (R_f 0.0), at which point the reaction was cooled to RT and the phases separated. The aqueous layer was re-extracted with DCM (2 x 10 mL). The combined organic layers were washed with brine (20 mL), dried ($MgSO_4$), filtered and the solvent removed *in vacuo* to afford the title product (185 mg, 90%) as a white amorphous solid; $[\alpha]_D^{24}$ +128.1 (c, 1.0 in $CHCl_3$); δ_H (500 MHz, $CDCl_3$), 2.00, 2.01, 2.02, 2.03, 2.04, 2.05, 2.07, 2.08, 2.12, 2.17, 2.19, 2.21, 2.22, 2.23 (66H, 14 x s, 22 x $COCH_3$), 2.27 (1H, d, $J_{1,SH}$ 9.8 Hz, SH), 3.76 (1H, dat, $J_{4,5}$ 9.7 Hz, J 3.5 Hz, H-5a), 3.92-4.08 (12H, m, H-4a, H-4b, H-4c, H-4d, H-4e, H-4f, H-5b, H-5c, H-5d, H-5e, H-5f, H-5g), 4.17-4.36, 4.49-4.56 (12H, m, H-6b, H-6b', H-6c, H-6c', H-6d, H-6d', H-6e, H-6e', H-6f, H-6f', H-6g, H-6g'), 4.39 (1H, dd, $J_{5,6}$ 3.6 Hz, $J_{6,6'}$ 12.2 Hz, H-6a), 4.48 (1H, dd, $J_{5,6}$ 3.2 Hz, $J_{6,6'}$ 12.3 Hz, H-6a), 4.62 (1H, at, J 9.5 Hz, H-1a), 4.73-4.78 (5H, m, H-2b, H-2c, H-2d, H-2e, H-2f), 4.82 (1H, at, J 9.5 Hz, H-2a), 4.88 (1H, dd, $J_{1,2}$ 4.0 Hz, $J_{2,3}$ 10.4 Hz, H-2g), 5.09 (1H, at, J 9.9 Hz, H-4g), 5.27 (1H, at, J 9.1 Hz, H-3a), 5.30-5.44 (12H, m, -1b, H-1c, H-1d, H-1e, H-1f, H-1g, H-3b, H-3c, H-3d, H-3e, H-3f, H-3g).

Example 17: Preparation of SBLCys156-S-SePh

Single site modification was investigated using a model-cysteine-containing protein, serine protease subtilisin *Bacillus lentus* mutant S156C (SBLCys156). SBLCys156 (10 mg) was dissolved in degassed aqueous buffer solution (1 mL, 70 mM CHES, 5 mM MES, 2 mM $CaCl_2$, pH 9.5). PhSeBr (5 mg, 0.02 mmol) was dissolved in

acetonitrile (200 µL), of which 150 µL (40 eq) was added to the protein solution and placed on an end-over-end rotator. After 30 min, the absence of free thiol was shown by Ellman's analysis (G. L. Ellman, K. D. Courtney, V. Andres, R. M. Featherstone, *Biochem. Pharmacol*. 1961, 7, 88). The reaction was placed on an end-over-end rotator for a further 30 min, at which point the reaction mixture was loaded onto a PD10 Sephadex® G25 column and eluted with 70 mM HEPES, 2 mM $CaCl_2$, pH 7.0. The protein fraction was collected and dialysed (MWCO 12-14 KDa) against water, (1 x 4L for 1 h, 2 x 2L for 30 min), to afford SBLS156C-S-SePh; m/z (ES^+) found 26864 calcd. 26870.

Example 18: Preparation of SSβGCys344Cys432-(S-SePh)$_2$

Multiple site modifications were investigated using a mutant of the thermophilic β-glycosidase from the archeon *Sulfolobus solfataricus* containing two cysteine residues (SSβG-Cys344Cys432). SSβG-Cys344Cys432 (1 mg) was dissolved in aqueous buffer solution (1 mL, 70 mM CHES, 5 mM MES, 2 mM $CaCl_2$, pH 9.5). PhSeBr (2 mg, 0.02 mmol) was dissolved in acetonitrile (200 µL), of which 20 µL (74 eq) was added to the protein solution and placed on an end-over-end rotator. After 1 h the reaction mixture was loaded onto a PD10 Sephadex® G25 column and eluted with (70 mM HEPES, 2 mM $CaCl_2$, pH 7.0) to afford SSβGCys344Cys432-(S-SePh)$_2$; m/z (ES^+) found 57700 calcd. 57697.

Example 19: Representative protein glycosylation with sugar thiols

SBLCys156 (1 mg) was dissolved in aqueous buffer solution (1 mL, 70 mM CHES, 5 mM MES, 2 mM $CaCl_2$, pH 9.5). The sugar thiol dissolved in water was added to the protein solution in the stated quantities (see table

below for equivalents) and the mixture placed in an end-over-end rotator. After 1 h the reaction was analysed by mass spectrometry.

Results

Protein[1]	Thiol	Equiv of thiol	Conv. %	ESI-MS Found (theory)
SBLCys156	GlcSH	5	>95	26908 (26909)
SBLCys156	GalSH	5	>95	26908 (26909)
SBLCys156	GlcNAcSH	1	>95	26944 (26950)
SBLCys156	GlcGlcGlcSH	5	>95	27228 (27233)
SBLCys156	GlcGlcGlcGlcGlcGlcGlcSH	10	>95	27878 (27881)
SSβG-Cys344Cys432	GlcSH	60	>95	57760 (57775)

Conv. = conversion as determined by ESI-MS

[1] Activated by reaction with phenyl selenium bromide to give the corresponding protein-S-Se-Ph or protein-$(S\text{-}Se\text{-}Ph)_2$ compound prior to addition of the thiol.

The results in the above Table demonstrate that the method of the invention provides high percentage conversion to the desired products using as little as one equivalent of thiol compound. Furthermore, the results demonstrate that the method of the invention can be used for single and multiple site protein glycosylations. The three glycosylation sites in SBL-Cys156 and SSβGCys344Cys432 are found in very varying protein structures and environments with

different levels of exposure, illustrating the broad applicability of the method of the invention.

Example 20: Representative protein glycosylation of SBLCys156 using GlcGlcGlcGlcGlcGlcGlc-SH

1-Thio-2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,6-tri-*O*-acetyl-4-*O*-(2,3,4,6-tetra-*O*-acetyl-α-*O*-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-α-D-glucopyranosyl)-β-D-glucopyranose (15 mg, 0.007 mmol) and sodium methoxide (2 mg, 0.007 mmol) were added to a stirred solution of MeOH (2 ml). After 2 h, (petrol:EtOAc, 1:2) indicated the formation of a product (R_f 0.0) with the complete consumption of the starting material (R_f 0.2). The reaction was neutralised with the addition of Dowex®-50 ion exchange resin after which point the reaction was filtered and concentrated *in vacuo*. The crude 1-thio-β-D-maltoheptaose was taken up into water (5 mL) of which 300 µL (11 eq) was added to a solution of SBLCys156-S-SePh (1 mg) in 500 µL of aqueous buffer (70 mM CHES, 5 mM MES, 2 mM $CaCl_2$, pH 9.5). The resulting solution was placed on an end-over-end rotator. After 1 h the reaction mixture was loaded onto a PD10 Sephadex® G25 column and eluted with 70 mM HEPES, 2 mM $CaCl_2$, pH 7.0. The protein fraction was collected to afford GlcGlcGLcGlcGlcGlcGlc-SBLCys156; m/z (ES^+) found 27878 calcd. 27881.

Example 21: Enzymatic extension of SBLCys156-S-GlcNAc

GlcNAc-SBLCys156 (3 mg) was dissolved in 1 mL of MQ water. Phenylmethylsulfonyl fluoride (PMSF) was added (50 µL of a 100 mg/mL solution in acetonitrile; 500-fold excess). The reaction mixture was incubated at room temperature for 30 minutes and purified over a Sephadex®

G-25 (PD-10) desalting column. The purity of the deactivated protein was assessed by ESI-mass spectrometry (found: 27100, calc. 27104). The protein fraction was lyophilized and re-dissolved in 1.0 mL of 0.1M sodium cacodylate buffer (pH 7.52). $MnCl_2.4H_2O$ (3.2 mg; 16 μmol) and uridine diphosphate-galactose (UDP-galactose, 2.3 mg; 3.4 μmol, Kyowa Hakko; 30-fold excess) were added. Recombinant bovine β-1,4-galactosyltransferase from *Spodoptera Frugiperda* (EC 2.4.1.22, 100 mU, Calbiochem) was added and the reaction mixture was incubated at room temperature for 40 min to afford Gal β1-4GlcNAc-SBL-Cys156 (ESI-MS, found 27265, calc. 27266).

This Example demonstrates that glycosylated proteins prepared according to the method of the invention may be further modified by reaction with suitable carbohydrate modifying enzymes, for example glycosyltransferases such as β-1,4-galactosyltransferase which selectively forms the Galβ1,4GlcNAc linkage.

Claims

1. A method of chemically modifying a protein, peptide or amino acid comprising at least one thiol group, the method comprising converting said thiol group into a selenenylsulfide group.

2. A method according to claim 1, wherein the conversion is carried out by reacting the protein, peptide or amino acid comprising at least one thiol group with a compound of formula Ia or Ib:

Ia Ib

wherein:
L denotes a leaving group; and
R denotes an optionally substituted alkyl group, an optionally substituted phenyl group, an optionally substituted benzyl group, an optionally substituted pyridyl group or an optionally substituted naphthyl group, or R forms part of or is attached to a solid support.

3. A method according to claim 2, wherein R is phenyl.

4. A method according to claim 2, wherein the compound of formula Ia or Ib is PhSeBr.

5. A method according to any one of claims 1 to 4, further comprising reacting the selenenylsulfide group in the protein, peptide or amino acid with an organic compound containing a thiol group.

6. A method of chemically modifying a protein, peptide or amino acid comprising at least one selenenylsulfide group, the method comprising

reacting the protein, peptide or amino acid with an organic compound comprising a thiol group.

7. A method according to claim 5 or claim 6, wherein the organic compound is a carbohydrate compound.

8. A method according to claim 5 or claim 6, wherein the organic compound is a protein, peptide or amino acid.

9. A protein, peptide or amino acid comprising at least one selenenylsulfide group.

10. A protein, peptide or amino acid according to claim 9, wherein the selenenylsulfide group is a group of formula:

-S-Se-R,

wherein R denotes an optionally substituted alkyl group, an optionally substituted phenyl group, an optionally substituted benzyl group, an optionally substituted pyridyl group or an optionally substituted naphthyl group.

11. A protein, peptide or amino acid comprising at least one selenenylsulfide group which is obtainable by the method of any one of claims 1 to 4.

12. A protein, peptide or amino acid comprising at least one disulfide bond which is obtainable by the method of any one of claims 5 to 8.